

Review of Cognitive Diagnostic Researches on English Reading

Huang Hongyan Duan Huiqiong*

School of Foreign Languages, Nanchang Hangkong University, Nanchang

Abstract: As the core of the new generation of test theory, cognitive diagnosis is a new diagnostic technique based on the development of modern cognitive psychology and psychometrics, modern statistical mathematics and computer science. It has been applied in the field of English education, for the English reading test based on cognitive diagnosis evaluation can measure students' knowledge structure, processing skills or cognitive process, its results can provide cognitive diagnostic feedback. At present, the cognitive diagnostic evaluation of English reading has become a research focus of foreign scholars and has also attracted wide attention from Chinese scholars. This paper reviews cognitive diagnosis from five aspects: its development, theoretical basis, its application in English reading and the difficulties and shortcomings of existing research, as well as its summary and prospect in order to promote the application of cognitive diagnosis in English education, especially in English reading tests in China.

Key words: Cognitive diagnosis; English reading; Review

Received: 2020-03-04 ; Accepted: 2020-03-24 ; Published: 2020-03-28

基于认知诊断的英语阅读研究 综述

黄洪燕 段惠琼 *

南昌航空大学外国语学院, 南昌

邮箱: englishduan2011@163.com

摘 要: 认知诊断作为新一代检测理论的核心, 是现代认知心理学, 心理测量学, 现代统计数学和计算机科学发展下的一种新型诊断技术。在英语教育领域, 认知诊断也得到了应用, 因为基于认知诊断评价的英语阅读测试可以测量学生的知识结构、加工技能或认知过程, 其结果可以在认知上提供诊断反馈。现在, 英语阅读的认知诊断性评价已成为国内外学者的主要研究热点, 引起了广泛关注。本文从认知诊断的发展概况、理论基础、在英语阅读中的应用以及现有研究的难点和不足、总结这五个方面对认知诊断进行了回顾和展望, 希望能促进认知诊断在中国英语教育中的应用。

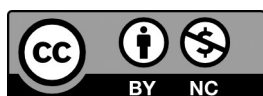
关键词: 认知诊断; 英语阅读; 综述

收稿日期: 2020-03-04; 录用日期: 2020-03-24; 发表日期: 2020-03-28

Copyright © 2019 by author(s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



在我国英语阅读测试中,阅读障碍时有发生,影响到英语学习者的学习热情和学习效果,怎样有效地预防和减少这类问题的存在是重要且必要的。认知诊断理论作为新一代心理和教育测试理论的核心,已成为当今国外学者研究的热点,并引起了中国学者越来越多的关注。在过去十年左右的时间里,中国还发表了一些与认知诊断相关的研究成果。2012年9月,第十届海峡两岸心理与教育测验学术研讨会在江西省南昌市举行。同时,出版了专著《认知诊断理论、方法与应用》,极大地推动了我国认知诊断的研究与发展。2018年6月7日,第一届语言学认知诊断研究国际研讨会在上海外国语大学举行。2018年6月23日,江西师范大学举办了智慧评估与学习学术研讨会。在CNKI全部文献中输入“英语阅读”和“认知诊断”两个关键词,获得5个结果表明,国内对英语阅读的认知诊断研究尚处于起步阶段。认知诊断通过获取被试在测试中可观察的反应模式,以推断其不可观察的知识状态,从而测量或评价个体特定的知识结构和处理技能。运用恰当的认知诊断模型诊断英语学习者在阅读过程里遇到的阅读障碍,从而考察他们的认知结构以及个体差异,可以有效地提高一些英语学习者在阅读过程中的阅读能力。本文在查阅现有文献的基础上,从认知诊断理论的发展概况、理论基础、在英语阅读中的应用、现有研究的难点和不足以及总结和展望这五个方面对认知诊断进行了回顾,以期促进我国英语教育领域的认知诊断研究。

1 认知诊断的发展概况

1.1 认知诊断的源起

认知诊断是指对个体知识结构、加工技能或认知过程(统称为属性, attribute)的诊断评估。有关认知诊断的研究始于20世纪70年代。依据Mislevy所言,心理测量学和教育测量学作为独立的学科分支,已经经历了一百多年的发展[1]。迄今为止,整个统计测验理论的发展历程大体可以分为以下两个时

期：第一时期是标准测验理论时期（Standard Test Theory），出现了一批包括经典测量理论（CTT）、项目反应理论（IRT）以及概化理论（GT）等在内的理论；第二时期是核心内容为认知诊断（Cognitive Diagnosis）的新一代测验理论时期（a New Generation of Test）。标准测验理论阶段中无论是经典测验理论还是项目反应理论都把被试的心理特征看作一个纯统计结构，在宏观上对个体进行整体的评价 [2]。然而，按照被试对题目所做出的反应来评估其能力这一做法，无法详细阐述能力所代表的结构。测试时只关注被试作答的结果，而忽视对作答过程的考察，这在一定程度上具有相当大的局限性。Mislevy 将这种仅强调对被试宏观能力水平评估的范式称为能力水平研究范式（Ability level paradigm） [3]。随着科学技术研究和社会生活的发展，人们对只得到个体宏观层面的评价感到越来越不满足，他们希望测验能深入到作答的认知加工过程当中，从微观层面揭示不同个体的认知加工水平如何，其特征又是怎样的。在传统测试中，总分相同或行为表现相同的个体不一定具有相同的心理加工过程，他们可能有不同的知识结构，解决问题的方法策略也迥然不同，而仅仅只拿到一个笼统的总得分是无法看出这些差异的。而新一代的测试理论克服了这一局限性，它于 1993 年由 Frederiksen 等人在著作《新一代测验的测验理论》（Test Theory for a New Generation of Tests）中正式提出 [4]。新一代的测试理论强调测试应该结合认知心理学的理论结果与现代统计学的方法，强调要同时从能力水平和认知水平两个方面进行测试，故而可以同时检测受试者的宏观和微观认知属性的认知诊断测试就被其视为核心内容 [2]。广义上而言，认知诊断是指建立起被试的分数和其内部的认知特征之间的关系，从而对其进行观察诊断。狭义的认知诊断则是指在教育教学领域中，根据被试对测试所测的技能和特征的掌握程度而对其进行分类。在测试领域，认知水平的研究是能力水平研究的深化和拓展。研究能力水平的目的是定位能力，而研究认知水平的目的是诊断认知能力的结构和特征 [5]。

1.2 认知诊断的理论基础

认知诊断是以认知心理学和心理测量学、现代统计数学和计算机科学三大

理论的发展为基础的。

1.2.1 认知心理学

认知心理学是一个理论体系，它指出，人类的精神和身体活动可以用计算机的信息处理来解释，并试图通过关注个体如何感知、记忆、思考和学习来调查思维是如何运作的。作为一门新兴的学科，它融合了人工智能、认识论、语言学、数学、神经科学和哲学等领域的思想，研究人的心理活动和心理加工过程并建立相应的模型。它可以为认知诊断测验的编制提供心理学方面的理论支持，用以指导项目的开发。将认知心理学应用于认知诊断研究就是要运用认知心理学的研究方法，对认知诊断测验当中所涉及的各种认知变量——诸如知识、技能、策略、加工过程和组成要素等进行认知分析，从而得到研究中所需要的心理模型。它不仅能明确被试进行正确作答所需的知识、技能、解决问题的策略和加工处理过程，而且还能明确项目特征、刺激条件和作答反应这三者之间的关系，从而提高测试编制过程中对难度等性能的预控性。因此，认知心理学是测验实现认知诊断功能的前提条件，也是认知诊断实施的基础和核心 [6]。

1.2.2 心理测量学

心理测量学一般而言指的是致力于测试、测量、评估等相关活动的心理学和教育的研究领域。该领域涉及对技能和知识、能力、态度、个性特征和教育成就的客观衡量。一些心理测量学研究者专注于评估工具的构建和验证，如问卷、测试、评分者的判断和个性测试。其他人则专注于与测量理论相关的研究，如项目反应理论（item response theory, IRT）；组内相关性。而心理测量学也是实现认知诊断的必要工具。仅仅通过认知分析获得一些理论依据是无法最终实现对被试个体差异的分析和解释的，测验的编制同样很重要。根据被试的实质性心理模型选择、编制项目和测试，从而得以详细描述被试之间的差异。Embretson 测试设计的核心是通过项目设计触发被试内在认知特征的外部表现，从而实现对内部认知特征的判断 [7]。在整个试验设计完成以后，还需要用测量数据来对理性推测得出的成果进行验证。具体操作是在测量模型中融合各种认知变量，运用现代统计学的方法来揭示个体的内部认知特征。

1.2.3 现代统计数学和计算机科学

基于项目反应理论和其他测量技术的认知诊断模型通常比较复杂,在一定程度上,已经达到了相当高的数学形式化程度。因此,为了深入分析实验数据,估计模型中的未知参数,识别认知诊断的模式,就需要具备统计数学和计算机科学的相关知识和技术。不若如此,认知诊断模型就仅仅只能停留在理论研究阶段,而不能真正应用于实际情况当中,其发展空间也会缩小,仅囿于理论层面了。

1.3 认知诊断的特征

Corter 指出,能实现诊断功能的新一代测验和传统测验有以下不同的特征 [8]。第一,传统测验只能测试出广泛的能力,而认知诊断测试测量的是特定的心理特征。传统的测试把被测对象的潜在特质 (latent trait) 看作一个“统计结构”,而认知诊断测试则试图分解这种“统计结构”,更深入地探索被测对象的内部心理特征。第二,传统测验假设不同层次的技能连续性地发生变化,测试的目的是在连续性标尺上对所测能力进行定位。而认知诊断测验测量的特质结果要么是“有”,要么是“无”。换句话说,认知诊断测试所测量的策略和技能,在测试结果上只会呈现有或无的状态。第三,在传统测验中,技能被认为要么是一维的,要么是多维的。而在认知诊断测验中,技能的组成成分可以与很多模式相联系,进行考察分析。主体在不同的情境中可能采用不同的方法,其技能的构成与时间或逻辑有一定的关联。还有,传统的测试采用项目一致性的几何模型,而在认知诊断测试中,因为认知诊断所测量的特质之间不光有量的差异,还有质的差异,所以使用的是项目一致性的离散性特征模型。

一般来说,认知诊断具有以下特点:第一,认知诊断以被试的个体差异为依据,对其进行评估诊断,同时调试教师的教学方案。其次,认知诊断的目的在于考察被试在测试中表现出的认知结构和解题程序。第三,认知诊断结果的解释,与描述被试在该学科领域所表现的心理语言模式有关 [9]。认知诊断测验不光是参照整个团体来解释测试的成绩,还从被试的作答反应结果提供他们的学习掌握度、优势和缺陷的具体诊断信息,为教师进行个别辅导提供了依据,

从而有针对性的开展矫正活动。认知诊断把认知心理学的理论成果与现代统计学方法相结合,开发了许多认知诊断模型,在模型中引入认知变量,对个体的心理机制和微观过程进行深入的考察和分析,为下一步的个体矫正训练提供了指导依据。

1.4 认知诊断在国内外的发展

我国学者将注意力主要放在教育领域的认知诊断,也就是狭义的认知诊断研究领域。要想开展认知诊断评估(Cognitive Diagnostic Assessment, CDA),需要进行一系列的系统化工作,包括定义认知模型、编制诊断测试、估计诊断模型的参数、报告诊断结果和制定有针对性的补救措施。而诊断测试的编制过程十分复杂:首先要确定诊断目标,然后根据诊断目标来确定认知属性、属性间的阶层关系,再对阶层关系加以验证,最后通过确定使用的认知模型来编制认知诊断测验并进行试测。这种评价过程与理论阶段的标准测试相比是一项非常困难的专业工作。它主要体现在难以准确确定认知模型、认知诊断测试的设计与编制过程复杂、较难实现认知诊断模型的参数估计以及诊断评估过程需要多科专家的参与等方面[2]。正因如此,我国的认知诊断研究大多都还处于起步的阶段,有一些理论研究,但在实际应用领域少有大规模的开展。通过认知诊断这一途径,人们可以更加深入地了解个体内在心理活动的规律及其加工机制,了解个体的知识结构、运用的认知策略和对属性的掌握情况,从而诊断出个体认知的强项和弱项。这将有助于被试自己和老师对其知识掌握情况有全面的了解,有助于被试进行有效的矫正性学习,有利于教师进行有针对性的教学。因此,认知诊断已成为当前国际心理测量学和教育测量学研究的一个重要热点。

认知诊断研究在国外的发展早于中国。最早可追溯到二十世纪六七十年代,认知心理学与心理测量学的结合开始受到研究者的关注。80年代,进一步开展了基础研究,并开始发表一些关于实际应用方面的研究成果。20世纪90年代,认知诊断的结果逐渐丰富,且已出版认知诊断评估专集。21世纪至今,认知诊断研究快速发展,吸引了越来越多的研究者的目光。目前已经发表了大量的学术论文,并开发了100多个认知诊断模型,研究者们也曾召开过几次以认知诊

断为主题的国际性会议 [5]。在教育实践过程中,英美等发达国家对认知诊断在教育改革中做出的贡献给予了相当的重视。2001年,美国正式通过了法案《不让一个孩子掉队》(NCLB or Public Law: 107-110),要求教育有针对性评估每个孩子的学习过程,不能仅仅给出简单的分数,还要提供详细的诊断信息。在台湾,郭伯臣带领的研究小组已开发了一种结合认知诊断的计算机化自适应测验,并在教育实践中进行了实验性应用。2003,香港的文剑冰博士开展了结合认知诊断的计算机化自适应测验研究 [10]。大陆的王家祺、刘红云利用融合模型诊断了小学生数学知识属性的掌握情况 [11]。张玉美、罗少茜采用认知诊断模型对八年级学生英语阅读能力的优劣势进行了分析和信息反馈 [12]。总体而言,中国在这方面的研究还处于起步阶段。但近年来,由于国家社科基金立项从2012年开始,已经有4项认知诊断相关的语言学立项,促进了此类研究的快速发展,研究成果逐渐增多。

2 认知诊断在英语阅读测试中的应用与研究结果

2.1 英语阅读测试中认知诊断的应用

目前在英语语言测试领域,认知诊断的研究以诊断听力和阅读能力为主,并逐渐进入对写作的诊断研究。就阅读而言,最早的关于英语阅读测试的认知诊断研究可以追溯到上世纪90年代。Sheehen曾于1993年在美国教育考试服务中心(Educational Testing Service, ETS)发表关于应用规则空间模型来诊断美国青少年文本加工技能的报告。随后,Buck应用规则空间模型诊断了托业(TOEIC)考试阅读试题的认知属性 [13]。Kasai和同事对阅读理解多项选择题进行了认知分析研究,对托福考试阅读理解部分重新进行了认知属性分析 [14]。2009年,Jang应用融合模型对设计用于学生的学术英语技能、帮助新托福考试的系统锋利英语(LanguEdge)考试阅读试题的认知属性进行了认知诊断 [15]。同年, Lee和Sawaki应用GDM模型、融合模型以及潜在分类模型分别对托福考试的阅读和听力试题的认知属性进行了认知诊断 [16]。到了2015年, Kim用

融合模型对成人二语分班阅读考试进行了认知诊断研究。然而, 中国对英语阅读测试的认知诊断研究还不是很多 [17]。蔡艳等人运用属性层次模型对中学生阅读能力的认知属性进行诊断 [18]。赵冠芳等人分别应用了 G-DINA 模型框架中的补偿模型和饱和模型对国际学生评估项目 (Programme for International Student Assessment, PISA) 阅读测试进行了认知诊断 [19]。李瑞、吴茜媛运用认知诊断测量技术及数字化技术开发了以英语阅读为切入的动态诊断评估与学习策略干预指导为一体的自主学习数字化在线应用系统 [20]。杜文博、马晓梅将认知诊断测评技术与树回归分析相结合, 以结合量化与质化的方法对英语阅读认知诊断测评模型的构建进行研究 [21]。张玉美、罗少茜采用认知诊断模型对八年级学生英语阅读能力的优劣势进行了分析和信息反馈 [12]。

综上所述, 英语阅读测试的认知诊断是目前认知诊断研究的重要课题。认知诊断技术的发展促进了英语阅读测试和阅读技巧的研究。反过来, 英语阅读测试和技能的研究又促进了关于新型认知诊断模型的开发。

2.2 英语阅读认知诊断测验的研究结果

到目前为止, 已有一批关于英语阅读认知诊断研究的文章发表了出来。不同的研究者从不同的角度出发, 界定出了迥异的认知属性。以下是从研究方法和技能两个方面对研究成果进行大致的梳理。

2.2.1 Kasai 的研究结果

结合英语阅读理解的加工模型和阅读理解多项选择题难度的影响变量这两个方面的研究结果, Kasai 在自己和同事对阅读理解多项选择题认知分析的研究基础上对托福考试中的阅读理解部分重新进行认知属性的分析, 并识别出了 16 个认知属性, 具体如下:

A1 低频词汇: 指的是整篇文章中包含较多的低频词汇, 考查个体对低频词汇的掌握度。

A2 信息位置明确指出: 指的是问题清楚地提到了文章的某个特定部分, 通过诸如“在第 2 行中、‘它’这个词指的是”这样的短语表示, 考查的是浏览

文章特定部分的技能。

A3 词汇重叠表明信息位置：指的是通过问题和文章的相关部分之间的词汇重叠，可以直接找到回答问题所需的必要信息所在的段落部分，考查的是浏览文章而后找出题目中的目标词汇的能力。

A4 信息位置不明显：指文章中回答问题的必要信息所在的部分不容易找到；没有明显的词汇重叠；问题包含对文章相关部分的复述性表达。考查的是识别语义等价的表达、理解比喻性语言以及辨别推断的技能。

A5 低频词汇：指正确选项中含有低频词汇，考核掌握低频词汇的能力。

A6 词汇重叠：指题目的正确选项与文章相关部分有词汇重叠的现象，考查的是利用相似词汇来获取正确答案的技能。

A7 超出文章的范畴：为了做出正确的选择，不仅仅是依靠文章中明确提供的信息，还需要做出推断和概括、得出结论、理解作者的意图和目的等。考查的是推断、概括和总结的能力及理解作者的意图、目的的能力。

A8 合理的错误选项的干扰：指的是错误选项看起来似乎合情合理，但事实上对于正确选项而言意思太过宽泛或太狭隘，考查的是准确理解文章和问题并排除错误选项的技能。

A9 理解句与句之间的关系：指通过利用衔接手段的知识来获得正确的选项，比如使用指称形式（例如代词、指示定冠词）、关键词的重复、替换和转换。考查的是利用各种衔接手段来理解句子之间的关系的的能力。

A10 修辞组织结构的知识掌握：指回答问题需要用到修辞组织的知识，包括论文陈述、主题句、文章思想的组织和信息流等。考查的是对文章组织结构的了解并运用这些信息和知识获得正确答案的能力。

A11 时间限制：指的是问题的答案出现在文章的最后三分之一部分，考查的是在规定的时间内迅速阅读完整篇文章并回答所有问题的能力。

A12 错误选项的词汇重叠：指题目的错误选项与文章相关部分有词汇重叠的现象，考查的是准确理解文章和问题，排除看似正确的选项的能力。

A13 复合句结构：指的是问题或目标句是复合句，由一个从属连词（比如“if”“when”“as”等）或者是关系名词、关系副词连接两个以上的从句，考

查的是处理语义复杂的长难句和理解不同部分之间关系的能力。

A14 不常见的句子结构：指问题或目标句中包含不常见的句子结构，如主语和动词的倒装，考查的是辨别不常见句型和理解句意的能力。

A15 利用选项来获取正确选项：指的是运用排除法一个一个进行选项的选择，考查的是回溯和使用选项来获得正确答案以及在不确定某一选项是否正确的前提下正确作答的能力。

A16 长的正确选项：指正确选项句子长于错误选项句，考查的是“对正确选项有时比错误选项句子要长”这一事实的了解。

Kasai 使用回归分析得出的回归方程的决定系数为 0.91，表明 16 个属性基本贴合项目的难度。不过，Kasai 还考虑到属性间的交互作用，通过实验证明了这 16 个属性还未能全部描述该测试所用到的重要属性，在进行认知诊断时还要考虑属性间的交互作用。

2.2.2 Jang 的研究结果

通过有声思维测验、问卷调查、课堂观察等，Jang（2005）采用统计测验及项目分析法在 NG TOFEL（Next Generation TOFEL）考试中确定了九种基本的阅读理解技能，它们分别是：

A1 CDV（Deduce word meaning from context），通过内容线索推断一个词语或短语的意义。

A2 CIV（Determine word meaning out of context），依靠被试的背景知识推断词意。

A3 SSL（Integrate syntactic and semantic pieces），利用词汇、语法等知识理解句子之间的关系与意义。

A4 TEI（Processing of explicit information），快速阅读文本某一部分，理解其意思。

A5 TIM（Paraphrasing implicit information），有选择的阅读文本的某个或数个段落以确认段落的意义。

A6 INF（Infer arguments from implicit information），略读文本，推断文意。

A7 NEG (Processing negation) , 对相关信息进行定位, 判断问题或选项中的信息是否正确。

A8 SUM (Summarize major ideas) , 归纳文本的主旨大意。

A9 MCF(Mapping contrasting ideas), 识别文章的主要论点, 并构建文本框架。

Jang 将融合模型用于认知诊断, 并成功地利用这九个属性对 2700 多名学生和教师的阅读能力进行了诊断。

2.2.3 Wang, Gierl &Leighton 等人的研究结果

Wang, Gierl & Leighton 通过调查有关第二语言英语的阅读理解文献 [22] , 识别出了八个认知属性, 它们分别是:

A1 BA 基础知识。

A2 US 对句子的内容、结构及功能的理解。

A3 UT 对文章片段的内容、结构及功能的理解 A4 PGS 分析作者的目的及表达策略。

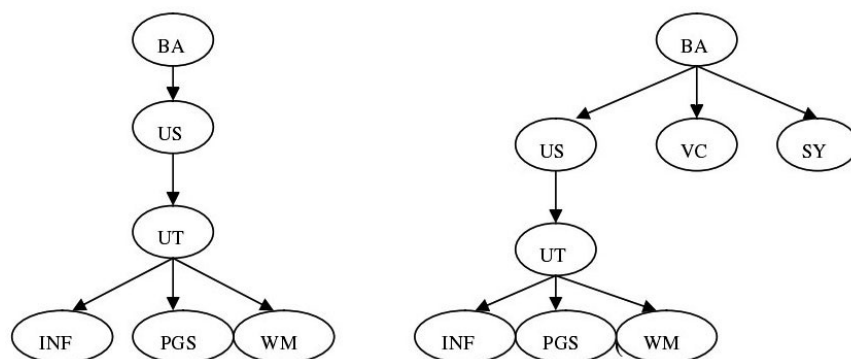
A5 WM 依据上下文推断词义。

A6 INF 依据背景知识进行推断。

A7 VC 对含有难词的文章段落的理解。

A8 SY 对复杂语句的理解。

此外, 他们通过属性之间的关系确定了两个层次的属性关系 (见下图) 。他们对数据进行分析诊断后证实第二种所表示的属性阶层关系在实际情况中更为贴合学习者的认知加工过程。



2.2.4 蔡艳的研究结果

蔡艳(2010)结合国外英语阅读理解的相关研究成果及教育部出台的《普通高中英语课程标准》和 NMET (National Matriculation English Test) 阅读理解测试要求,运用分析法对我国高考阅读理解试题进行了认知属性界定[23],具体如下:

A1 VC 词汇:对问题、选项或目标句中出现的较难词汇的理解,对应编码加工。

A2 CDC 复句理解:对长而语义复杂的句子以及不常见句型的加工及理解能力,对应编码加工与一致性加工。

A3 CRC 理解句子间的关系:利用各种一致性知识理解句子间关系的能力,对应一致性加工。

A4 RSC 修辞组织结构的理解:正确回答问题所需要的组织结构和修辞的知识,对应编码加工与一致性加工。

A5 INF 推断:根据文章信息及背景知识对词意、作者的目的等进行推断的能力。

A6 MNI 信息匹配:浏览和匹配解决问题所需要的文章信息的能力,对应文本映射加工与编码加工。

A7 正确选项的加工水平:正确理解问题及文章信息并准确排除错误选项的能力,对应编码、一致性加工与评估反应选项的真实性加工。

A8 Summary 主旨大意:对给定信息的理解能力和总结归纳能力,对应文本映射加工与编码加工。

2.2.5 陈慧麟,赵冠芳的研究结果

陈慧麟,赵冠芳应用 LLM 认知诊断模型对 500 名考生的 PISA 英语阅读测试结果进行实证分析[19],确定了 5 个阅读理解的认知属性,它们分别是:

A1 直接信息的提取——在篇章中查找识别显性信息。

A2 内容的总体把握——把握篇章中心思想及上下文的归。

A3 概念及关系的阐释——阐释复杂概念及复杂关系。

A4 间接信息的推测——根据言外之意推测隐性信息。

A5 篇章评价——评价篇章文体及作者意图。

他们具体采用的方法：首先，选取8位母语是英语的中学生完成英语阅读理解测试，再进行有声思维测试，要求学生根据自己做题时的感受从现有阅读技能分类中选取各题所测的阅读认知属性。然后，邀请6位语言学和教育学领域的大学教师进行专家论证，最终确定出这五个包含了PISA阅读试题测试规范规定要考查的阅读技能的认知属性。

2.2.6 马晓梅、杜文博的研究成果

马晓梅和杜文博（2018）通过文献回顾、专家确定和有声思维，根据文本阅读双加工理论和“建构—整合”模型，共同确定了8个阅读认知属性的定义，具体如下：

A1 理解句子字面意思，即运用词汇、语法等知识理解句子的结构及字面意思。

A2 理解篇章字面意思，指的是整合多个句子的字面意思来理解句子间的关系以及篇章表达的字面意思。

A3 推断词意，即根据上下文推断词意。

A4 连贯性推理，即根据上下文进行推理以填补阅读过程中的句子前后信息的空隙。

A5 精加工推理，即通过背景知识和文本信息对文中句子隐含意义、因果关系、作者目的和意图等进行推断。

A6 主旨大意，即由多个句子信息或文章部分内容概括出文章主旨、主题。

A7 定位相关信息，即根据问题中提到的信息，浏览全文寻找相关细节，进行信息定位。

A8 排除错误选项，即根据段落或篇章含义，从各个选项中对错误选项进行逐一排除，最终得出正确答案。

综上所述，基于英语阅读的认知诊断测验的研究已经取得了一些成果。对阅读理解进行诊断性研究是可行的，但比较困难，因为不同的研究者使用不同的方法分析相同的测试可能会得出不同结果，即便是同一研究者对同一份测验进行分析，也可能因认识的程度不同而得到不一致的结果。故可知，对英语阅读的认知诊断研究还有很长的路要走。

3 现有研究的难点与不足之处

3.1 现有研究的难点

将认知诊断应用于英语阅读的研究始于20世纪90年代,至今已有近30年。然而,到目前为止,国内外对英语阅读的认知诊断研究还大多停留在理论研究和小规模应用领域,在大规模社会考试中的应用还很少。在我国,英语阅读诊断测验的研究还属于起步阶段,发表于学术刊物的论文数量并不多。本文从国内外英语阅读认知诊断的现状出发,认为认知诊断为英语阅读测试提供了良好的研究方向,其目标明确,思路清晰,其结果也具有重大的现实意义。但在英语阅读的认知诊断研究中要想得到理想的研究结果却有一定的难度,这大概是至今为止还较少地应用于实际的原因所在。通过上面的分析不难看出,研究难点与理论基础有很大关系。首先,从心理测量学的角度来看,很难建立一种完全合适的、实用的认知模型。构建的模型越复杂就能越好地容纳认知变量和认知心理研究的结果,所获得的参数就能更准确地解释个体的心理结构特征,但在此过程中参数评价等技术难点一直难以得到解决。而使用简单模型的话,参数评估技术的确变得容易解决,可同时所得的参数便不足以很好地解释心理特质了。其次,认知诊断分析工作需要研究人员有丰富的认知心理学知识、理论和实践经验。对于不是这个领域的研究人员来说,认知心理学的知识和术语等显得十分陌生,不仅要接受,而且要熟练地使用它,就需要更长期的学习。此外,仅凭认知模型还不足以分析和解释个体差异。为了更深入地分析实验数据,估计模型中的未知参数,需要具备统计学和计算机科学的相关知识和技术支持。研究人员要重视测验的设计,即根据被测量对象的心理模型来选择、编制测验项目,进行测试。在设计好测试后,需要使用适当的测量模型,使理论层面的想法获得实证性的数据。认知诊断测试需要较大的测验规模,而且不少模型公式复杂,计算繁琐,数据量大,这些都是一直以来研究的难点。

3.2 现有研究的不足

尽管现如今已经有诸多学者将认知诊断应用于英语阅读理解测试当中,并

取得了较为丰富的成果,研究的内容也涉及到方方面面,但就中国国内而言,现有的研究还不能算多。实际上,国外对英语阅读理解的认知诊断研究起步较早,成果较多,而国内涉及该方面的研究还处在起步的阶段。在国外的研究中,属性都是依据特定的测验来界定的,这种方式有其合理性,但也有一定的局限性。以 Jang 和 Kasai 的研究为例,虽然他们都是基于托福考试来界定认知属性的,但界定的出发点并不一样。Jang 着重从考试所测试的能力和测试时需要使用的阅读方式这两个角度对认知属性进行识别,而 Kasai 则是从解答问题所涉及的文章信息及选项和选题策略的角度来识别属性。不过, Jang 和 Kasai 对某些属性的定义存在冗余的现象。比如, Kasai 研究中提出的 A1 和 A5 属性都是考查个体的词汇量,属性 A3 和 A6 都是考查个体利用重叠词汇来查询与解题相关的文章信息的能力,如此一来,就将认知模型复杂化了。除此以外,国外学者们识别出的某些属性与中国当前的国情不相符合。比如,在我国的测试中,对 Jang 所提出的 TEI (快速阅读文本某一部分,理解其意思) 属性和 MCF (识别文章的主要论点,并构建文本框架) 属性少有要求。Kasai 提出的 A16 属性考核的是学生个体是否了解‘正确选项有时比错误选项句子要长’这一事实,可是在中国测试中并不存在对这一属性的考查。又比如, Wang 虽然是基于中国的大学英语测试来界定认知属性的,但在这一过程中忽略了项目本身的特点,比如说,错误选项看起来越是合情合理,那么该项目的难度就越大。综上可知,在研究有关英语阅读理解的认知诊断过程当中,结合我国国情、教学大纲和考试大纲是十分有必要的。只有这样,才能使研究更好地服务于我国的英语教学及个体的知识获取,具备更大的实用性。

4 总结

认知诊断是新一代测试理论的核心理论,从实用功能的角度看,它已经走出了只注重结果而忽视过程,只注重结果而忽视补救的传统测试领域。认知诊断理论是从掌握学生的认知特征和属性入手的,它为被试的学习缺陷提供了良好的诊断方法。在此基础上,对其不足与欠缺之处进行弥补性的教学,这无疑是认知诊断的一大亮点。基于英语阅读的认知诊断为获取被试掌握了哪些知识,未掌握什么知识等信息提供了有力的支持。并可由此提出补救措施,以防止阅

读理解中所犯的错误不断重现,从而提高教学效率,这是我国研究人员未来的研究方向。基于本研究的各种困难,对英语阅读的认知诊断感兴趣的研究者需要钻研现有的认知诊断理论和方法,补充认知心理学和统计测量学、统计数学和计算机科学等必要的知识,认真探索每一环节的工作。英语阅读的认知诊断研究应吸引更多相关领域的研究者,形成分工协作的研究队伍,从而有效地促进我国英语阅读的认知诊断研究。

当然,认知诊断可应用的语言测试领域远不止阅读,还有听力、写作、口语等,这些都是教育者应研究的方向和努力的目标。认知诊断可以科学地描述测试中所包含的认知属性,并且精确地诊断出语言的技能结构。其在语言测试中的应用不仅可以促进语言及语言测试的理论研究,还对改进语言测试的设计和编制,改善语言教学的内容及方式,建立精确的语言技能评分机制,推广个性化语言教学起到了促进的作用,是必要且可行的。

基金项目

本文为江西省教改课题“大学英语阅读障碍的认知诊断及补救教学”(编号:JXJG-13-8-19)的部分成果。

参考文献

- [1] Mislevy R J. Foundations of a new test theory [M] //N Frederiksen, et al. (Eds.), Test theory for a new generation of tests. Hillsdale, NJ: LEA, 1993: 19-39.
- [2] 谢美华. 初中生现代文阅读理解能力的认知诊断评估研究[D]. 江西师范大学, 2014.
- [3] 戴海琦, 谢美华, 丁树良. 我国大陆认知诊断研究的文献计量分析[J]. 南京师大学报(社会科学版), 2013(6): 88-97.
- [4] Frederiksen N, et al. Test Theory for a New Generation of Tests [M]. Hillsdale, NJ: LEA, 1993: 19-39.
- [5] 刘声涛, 戴海琦, 周骏. 新一代测验理论—认知诊断理论的源起与特征

- [J]. 心理学探新, 2006, 26(4): 73-77.
<https://doi.org/10.4037/ccn2006.26.6.77>
- [6] 陈慧麟. 语言测试中的认知诊断及其应用流程[J]. 外语测试与教学, 2015(2): 51-60.
- [7] Embretson S E. Introduction to the Problem of Test Design [M] // Embretson S E (Eds.), Test design: developments in psychology and psychometrics. Harcourt Brace Jovanovich, Publishers, 1985: 3-17.
<https://doi.org/10.1016/B978-0-12-238180-5.50005-7>
- [8] Corter J E. Using clustering methods to explore the structure of diagnostic tests [M] // Paul D. Nichols, Susan F. Chipman, Robert L. Brennan (Eds.), Cognitively diagnostic assessment, Lawrence Erlbaum Associates, Inc, 1995: 307.
- [9] Shanon B. The representational and the representational: An essay on cognition and the study of mind [M]. New York: Harvester Wheatsheaf, 1993.
- [10] 文剑冰. 多维计算机化自适应测验的选题策略比较[C] // 中国心理学会. 第十二届全国心理学学术大会论文摘要集. 中国心理学会: 中国心理学会, 2009: 1.
- [11] 王家祺, 刘红云. 融合模型在小学数学认知诊断评价中的应用[J]. 心理学探新, 2012, 32(5): 423-428.
- [12] 张玉美, 罗少茜. 基于认知诊断模型的英语阅读测试诊断信息反馈[J]. 外语测试与教学, 2018(3): 55-60+64.
- [13] Buck G, Tatsuoka K, Kostin I. The Subskills of reading: Rule-space analysis of a multiple-choice test of second language reading comprehension [J]. Language Learning, 1997, 47(3): 423-466.
<https://doi.org/10.1111/0023-8333.00016>
- [14] Kasai. The rule space model applied to the reading comprehension section of the Test of English as a Foreign Language (TOEFL) [D]. University of Illinois at Urbana-Champaign, 1997.

- [15] Jang E E. Cognitive diagnostic assessment of L2 reading comprehension ability: Validity Arguments for Fusion Model application to LanguEdge assessment [J] . Language Testing, 2009, 26 (1) : 31-73.
<https://doi.org/10.1177/0265532208097336>
- [16] Lee Y, Sawaki Y. Application of three cognitive diagnosis models to ESL reading and listening assessments [J] . Language Assessment Quarterly, 2009, 6 (3) : 239-263. <https://doi.org/10.1080/15434300903079562>
- [17] Kim A H. Exploring ways to provide diagnostic feedback with an ESL placement test: Cognitive diagnostic assessment of L2 reading ability [J] . Language Testing, 2015, 32 (2) : 227-258.
<https://doi.org/10.1177/0265532214558457>
- [18] 蔡艳, 丁树良, 涂冬波. 英语阅读问题解决的认知诊断 [J] . 心理科学, 2011, 34 (2) : 272-277. <https://doi.org/10.25291/VR/34-VR-277>
- [19] 陈慧麟, 赵冠芳. 认知诊断的应用——语言测试研究的新阶段 [J] . 外语测试与教学, 2013 (2) : 1-9+31.
[https://doi.org/10.1016/S0264-410X\(13\)00434-9](https://doi.org/10.1016/S0264-410X(13)00434-9)
- [20] 李瑞, 吴茜媛. 基于数字化的英语阅读认知诊断模式构建及其在线应用研究 [J] . 海外英语, 2014 (23) : 1-2+4.
- [21] 杜文博, 马晓梅. 基于认知诊断评估的英语阅读诊断模型构建 [J] . 外语教学与研究, 2018, 50 (1) : 74-88+160-161.
- [22] Wang C, Gier M J, Leighton J P. Investigating the Cognitive Attributes Underlying Student Performance on a Foreign Language Reading Test: An Application of the Attribute Hierarchy Method [C] // Paper presented at the Paper prepared for the graduate student poster session at the 2006 annual meeting of the National Council on Measurement in Education, San Francisco, California, 2006.
- [23] 蔡艳. 群体水平的英语阅读问题解决能力评估及认知诊断 [D] . 江西师范大学, 2010.